

Abstract

Most European countries have regulatory requirements or guidelines for reverberation time in classrooms which have the goal of enhancing speech intelligibility and reducing noise levels in schools. At the same time, school teachers suffer frequently from voice problems due to high vocal load experienced at work. With the aim of improving working conditions for teachers, this article presents guidelines for classroom acoustics design that meet simultaneously criteria of vocal comfort and speech intelligibility, which may be of use in future discussions for updating regulatory requirements in classroom acoustics. Two room acoustic parameters are shown relevant for a speaker: the voice support ST_V , linked to vocal effort, and the decay time $DT_{40,ME}$ derived from an oral-binaural impulse response, linked to vocal comfort. Theoretical prediction models for room-averaged values of these parameters are combined with a model of speech intelligibility based on the useful-to-detrimental ratio U_{50} and empirical models of signal-to-noise ratio in classrooms in order to derive classroom acoustic guidelines, taking into account physical volume restrictions linked to the number of students present in a classroom. The recommended values of reverberation time in fully occupied classrooms for flexible teaching methods are between 0.45 s and 0.6 s (between 0.6 and 0.7 s in an unoccupied but furnished condition) for classrooms with less than 40 students and volumes below 210 m³. When designing larger classrooms, a dedicated acoustic study taking into account considerations about geometry, material and speaker/audience placements should be made, which can help to increase the voice support and reduce the vocal effort.

Speaker-oriented classroom acoustics design guidelines in the context of current regulations in European countries

David Pelegrín-García

david.pelegringarcia@fys.kuleuven.be

Jonas Brunskog

jbr@elektro.dtu.dk

Birgit Rasmussen

bir@sbi.aau.dk

January 16, 2014

1 Introduction

The acoustic design of school classrooms, both in terms of noise control and room acoustics, is relevant because it affects the quality of oral communication between teachers and students, which is still the most common way of teaching and learning. Students need a physical environment which preserves speech intelligibility, makes the oral message easy to listen and engages them into the process of learning. When this is achieved, the acoustic environment is no longer a barrier for students' personal development and academic achievement. At the same time, teachers' speech needs to be understood; they have to speak comfortably and without straining their voice. Two concepts related to the voice of the teacher are defined in this context: vocal comfort and vocal effort.

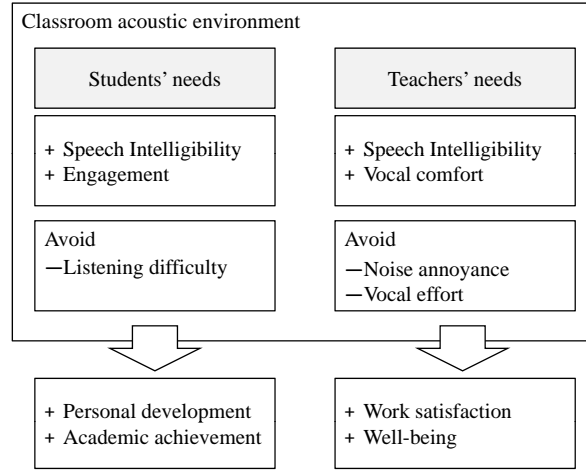


Figure 1: Needs of students and teachers in classrooms related to the acoustic properties of the environment

The *vocal comfort* in a room is a subjective attribute that is directly correlated to the positive evaluation of the room for speech purposes and to the perceived support, and negatively correlated to the feeling of having to raise the voice and to the tiredness after speaking longly in the room [1]. The vocal effort, according to Traunmüller and Eriks-son [2], is a physiological magnitude different from vocal intensity, which accounts for the changes in voice production required for the communication at different distances, under different noise or room acoustics conditions. Some descriptors of vocal effort are vocal intensity, fundamental frequency, phonation time, and spectral distribution. Moreover, the student-activity noise should be as low as possible to minimize teachers' discomfort. When these needs are met, teachers might experience higher satisfaction at work and have a good quality of life (with low stress, anxiety and vocal load). Figure 1 summarizes the needs of students and teachers regarding the acoustic environment. Figure 2 shows the most common sources of noise present in classrooms, such as noise from traffic, installations, neighboring activities, noise generated in corridors and play-grounds, and activity noise generated inside the classroom. The noise generated by the students inside the classroom is considered by teachers to be a stress factor [3, 4].

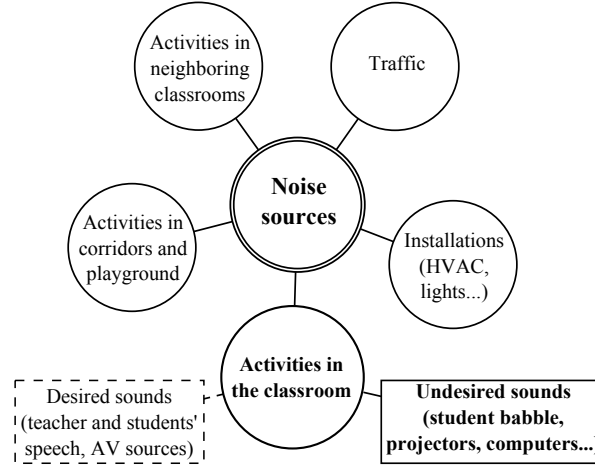


Figure 2: Noise sources commonly present in classrooms. Desired (non-noise) sound sources are the voices from teacher and students and the audio from audiovisual resources.

Excessive noise and late reverberation degrade speech intelligibility (see e.g. [5]). The effect of degradation is higher for younger students [6, 7] and those with impairments in hearing or developmental skills than for the general population. Early studies by Bradley [8] and Houtgast [9] found that speech intelligibility improved in adults for A-weighted Signal-To-Noise Ratios (SNR) up to +15 dB. Later, Bradley and Sato [6] suggested that the SNR that makes 75% of the students achieve 95% of speech intelligibility scores above grade 6 should be +15 dB, but for students in grade 1 the SNR should be +20 dB.

Most European countries have regulatory requirements or guidelines regarding the acoustic design of classrooms, with the goal of providing ideal conditions for teaching and learning in terms of speech intelligibility. These requirements or recommendations establish limit values of service-equipment and traffic noise and reverberation time (or *target* values in some countries) and minimum values of sound insulation.

Besides affecting speech intelligibility, noise and classroom acoustics affect the vocal comfort and vocal health of the teacher. The latter has recently been studied in a

joint Swedish-Danish project [10]. In a prevalence study of voice problems among teaching staff, 487 teachers of 22 schools in southern Sweden answered a questionnaire regarding voice problems [11]. The results showed that 13% of teachers suffered from voice problems at the time of the investigation. Room acoustics was pointed out as one of the causes leading to voice problems. Moreover, a link was found between the voice power level used by a talker in a room and the amplification that a room produces on his voice at his own ears, in the presence of low noise levels [12, 13]. This effect is related to the Lombard reflex (increase of voice power level of a talker in the presence of noise) [14] and the café effect [15] (which is related to the influence of the acoustic absorption in a room with several talkers on the voice power levels used by them; a decrease in absorption increases the energy of the reverberant field and thus the sound pressure levels, which in turn lead the talkers to raise their voices due to the Lombard effect). A model for the café effect in eating establishments has been proposed by Rindel [16].

After introducing a summary of some of the current European regulatory requirements on classroom acoustics, the present paper describes the room acoustics parameters relevant for a speaker, their connection to subjective attributes (vocal effort and vocal comfort), and prediction models for their spatial average, in terms of volume, surface area and statistical reverberation time. With these prediction models, some initial guidelines based on vocal comfort optimization are given. Simple models of speech intelligibility based on the early-late ratio C_{50} and the useful-to-detrimental ratio U_{50} , combined with empirical models of student-activity noise and voice power reported by Hodgson *et al* [17], are used with the aim of refining the initial guidelines and find classroom acoustic conditions which simultaneously meet criteria of speech intelligibility for students and vocal comfort for teachers.

Room acoustics parameters relevant for a speaker, their relationship with voice, and their prediction models have been presented separately in previous research [1, 13, 18]. Thus, the present article contributes to unify that scattered information, presenting

findings in a coherent manner, and to relate speaker-oriented requirements with speech intelligibility metrics in order to derive novel practical guidelines for classroom acoustics design. These guidelines are compared to current European regulations of reverberation times in classrooms, pointing out aspects to consider in future updates of these regulations.

2 Current classroom acoustics regulations in European countries

Classroom acoustics is a concern of increasing awareness, as demonstrates the increasing number of countries in Europe which have included classroom acoustic regulatory requirements in their building codes during the last 5-10 years (e.g. Spain) or which have made these requirements more strict (e.g. Denmark and Norway).

The present paper does not aim at making an exhaustive overview of the current regulations in all the European countries, but to represent a fair overview of the trends in regulatory requirements regarding reverberation time (other room acoustic parameters including maximum allowed indoor ambient noise levels and minimum sound insulation values, while relevant, are out of the scope of this paper). For this, six European countries with regulatory requirements for classroom acoustics have been chosen: Denmark, France, Germany, Norway, Spain, and United Kingdom (UK). The building regulations and the documents containing the requirements/guidelines for classroom acoustic design are summarized in Table 1. A similar comparison of regulatory requirements in Nordic countries can be found in [19]. All the countries in Table 1, with the exception of Germany, have national building codes that, in different ways, link to the requirements for classroom acoustics. The building legislation in Germany is competency of the *Länder* and thus, there is no building code common to the whole country. For specific constructions, the German jurisdiction establishes that recognized technical rules

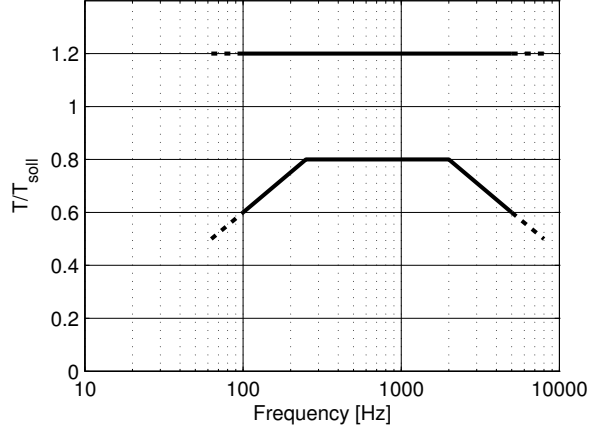


Figure 3: Maximum and minimum limits of reverberation time (relative to the target reverberation time T_{soll} shown in Table 2) in frequency bands set by DIN 18041:2004 [20].

(Anerkannte Regel der Technik) should be used. The standard DIN 18041:2004 [20], containing the classroom acoustics design guidelines, is a recognized technical rule.

A summary of the requirements regarding reverberation time in ordinary classrooms for general primary and secondary education (and thus excluding sports halls, music rooms, workshops and other specialized educational spaces) is shown in Table 2. Some of the countries (Denmark, Norway, UK, Spain) establish maximum values for the reverberation time whereas others (Germany, France) establish target ranges as a function of the volume. In particular, Germany establishes a relationship between the volume and the target reverberation time T_{soll} (*soll* is the German word for target) and defines a range of accepted values around that target reverberation time for different frequency bands (see Fig. 3). This is a trend that other European countries follow (e.g. Austria [32], Switzerland [33], Belgium [34]).

The UK proposes different maximum reverberation times for primary and secondary school classrooms, in response to the fact that speech degradation due to late reverberation is worse for younger students (primary education) than for elderly ones (secondary

Table 1: Summary of the norms containing the classroom acoustic design guidelines and the regulatory building codes.

Country	Building Code	Standard/Guideline	Year	Comments
Denmark	BR 2010 [21]	—	—	Requirements within Building Code. It refers to [22] for details.
France	CCH [23]	Arrêté du 25 avril 2003 [24]	2003	Application details for the requirements are described in [25].
Germany	(see comment)	DIN 18041:2004 [20]	2004	There is no common Building Code for schools to all federal states in Germany. DIN 18041:2004 is a recognized technical rule (Anerkannte Regel der Technik) required in the design of schools.
Norway	TEK'2010 [26]	NS8175:2012 [27]	2012	The Building Code expresses the requirements as Class C of the Classification Scheme NS8175:2012
Spain	CTE [28]	DB-HR [29]	2009	
UK	Building Regulations [30]	2010 BB93 [31]	2003	

Table 2: Regulatory requirements for classroom reverberation time in different European countries, in addition with details about the conditions of the criteria. All values are referred to unoccupied but furnished classrooms (unless otherwise expressed).

Country	Required T [s]	Freq. range [Hz]	Details of T
Denmark	≤ 0.6	125–4000	T_{20} according to ISO 3382-2. Max. in each 1/1 octave band. At 125 Hz: max. +20% accepted
France	$V < 250 \text{ m}^3$: $0.4 \leq T \leq 0.8$ $V > 250 \text{ m}^3$: $0.6 \leq T \leq 1.2$	Average 500–1000–2000	
Germany	$T_{\text{sol1}} = 0.32 \log \frac{V}{\text{m}^3} - 0.17$ ($V = 100 \text{m}^3 \rightarrow T_{\text{sol1}} = 0.47\text{s}$ $V = 250 \text{m}^3 \rightarrow T_{\text{sol1}} = 0.60\text{s}$)	100–5000	Occupied values. Unoccupied values +0.2 s. Limits for T as a function of freq. shown in Fig. 3. Values in classrooms for HI should be 20% lower.
Norway	≤ 0.5	125–4000	T_{20} according to ISO 3382-2. Max. in each 1/1 octave band. At 125 Hz: max. +30% accepted
Spain	$V < 350 \text{ m}^3$: ≤ 0.5	Average 500–1000–2000	Unfurnished: $T < 0.7\text{s}$
UK	Nursery & primary: ≤ 0.6 Secondary: ≤ 0.8 For HI: ≤ 0.4	Average 500–1000–2000	Values for unfurnished spaces. T may gradually increase at low frequencies (250 Hz and 125 Hz) but no more than 30% over limits

education) [6, 7]. In addition, they define maximum reverberation times for other educational spaces (e.g. music rooms, open-plan areas, sports halls) which are not within the scope of the present article.

The countries in Table 2 verify the compliance of the requirements in different ways. In all countries, calculations are requested during the design stage. Moreover, in Denmark and Norway, local authorities (when approving a new building project) can demand that *in situ* measurements are conducted. In the UK, the legislation recommends but does not require carrying out verification measurements in schools after they have been newly built.

The Norwegian Building Code expresses the acoustic requirements of classrooms as class C of Classification Scheme NS8175:2012 [27]. A Classification Scheme sets acoustic criteria (in terms of reverberation time, indoor ambient noise levels and sound insulation) to classify classrooms (and other spaces) in different levels of acoustical quality, called classes. In NS8175:2012 [27], class A has the highest acoustic quality and class D has the lowest (Classification Schemes in other countries might have different criteria and denomination for the classes). In NS8175:2012 [27], reverberation time in classrooms should be not higher than 0.4 s to achieve class A or B, 0.5 s for class C or 0.6 s for class D. Sweden [35] and Iceland [36] also have Classification Schemes. The trend in Classification Schemes is that classrooms are rated better the shorter the reverberation time (and also the lower the indoor ambient noise level and the higher the sound insulation).

3 Room acoustics parameters for a speaker

Traditional room acoustic parameters (e.g. reverberation time T_{60} , early-to-late ratio C_{50} , and speech transmission index STI) characterize the propagation of sound between a source and a receiver who are distant from each other. Instead, room acoustics

parameters that aim at characterizing the propagation of sound between the mouth and the ears of a speaker are derived from an *oral-binaural room impulse response* (OBRIR), i.e. an impulse response measured at a microphone located at the end of the ear canal of a dummy head when a loudspeaker inside its mouth acts as the source. The derivations that will follow assume the use of a head and torso simulator (HATS) from Brüel & Kjær (Brüel & Kjær Sound & Vibration Measurement A/S; Nærum, Denmark) type 4128. Two parameters are found relevant: the *voice support* ST_V and the *decay time* $DT_{40,ME}$.

3.1 Voice support

The voice support ST_V is a measure of the degree of amplification of a room to the voice of a speaker at his own ears [18]. More specifically, it is defined as the difference between the reflected sound level (L_R) and the airborne direct sound level (L_D) of the voice of a speaker, as found in an OBRIR.

$$ST_V = L_R - L_D \text{ (dB)}. \quad (1)$$

In practical terms, the airborne direct sound level is calculated by windowing the first 5 ms of an OBRIR whereas the reflected sound level is calculated from the later portion of the OBRIR (5 ms to ∞). For this calculation, it is assumed that there are no obstacles closer than 1 m to the measurement device. If there would be, the direct sound should be derived from calibrated measurements in an anechoic chamber, as in [12]. ST_V is minimum (ideally approaching $-\infty$ dB) in free-field, where there is no reflected sound, and becomes higher the smaller the room and the higher the reverberation time.

3.1.1 Relation with voice

The voice support is related to the *vocal effort* experienced by a speaker in different rooms [1, 12, 13, 37–39], and one of its manifestations is variations in the radiated voice power level. An empirical model that represents the average variations in voice power level ΔL_W of a speaker, with respect to the average voice power level used in free-field ($L_{W,0}$), as a function of ST_V [12, 13] is

$$\Delta L_W = L_W - L_{W,0} = \begin{cases} -13 - 0.78 ST_V \text{ (dB)}, & -14.5\text{dB} \leq ST_V \leq -6.5\text{dB} \\ 0.5 - 135 \log(10^{ST_V/10} + 1) \text{ (dB)}, & ST_V \leq -14.5\text{dB} \end{cases} \quad (2)$$

This equation is derived from measurements in a teaching-like setting with a silent audience. The variations of L_W in the presence of activity noise levels larger than 45 dB do not follow the model of Eq. (2), as the activity noise depends itself on the degree of amplification (or absorption) of the room. The threshold of 45 dB is found by Lazarus [40] as the minimum level that elicits the Lombard reflex in a speaker.

3.1.2 Prediction model

The prediction model for the average ST_V in a room [18] disregards the importance of the surroundings of the speaker in determining the actual ST_V at the speaker position and provides a unique value for a room, averaged across positions, assuming the systematic presence of the direct sound and a floor reflection; all other reflections are assumed diffuse. The model is expressed as

$$ST_V = 10 \log \left(\frac{4(1 - \bar{\alpha})}{S\bar{\alpha}} + \frac{Q^*}{4\pi(2d)^2} \right) + \Delta L_{\text{HRTF}} - K \text{ (dB)}, \quad (3)$$

where $\bar{\alpha}$ is the mean absorption coefficient of the room, S the total surface area, Q^* the directivity of a speaker in the downward direction, d the distance from the mouth to the floor ($= 1.5$ m), ΔL_{HRTF} the magnitude of the diffuse-field head-related transfer

function (HRTF, in dB), and K the difference between the SPL at the eardrum (in dB re $20\mu\text{Pa}$) and L_W (in dB re 1pW). This model contains the following terms:

- Diffuse-field attenuation of sound, indicated by the term $4(1 - \bar{\alpha})/(S\bar{\alpha})$ inside the logarithm. In the rest of the article $\bar{\alpha}$ will be derived from Eyring's formula ($\bar{\alpha} = 1 - \exp(-0.161V/(T_{60}S))$) rather than Sabine's ($\bar{\alpha} = 0.161V/(T_{60}S)$)—as opposed to the original reference [18]—, where V is the room volume and T_{60} the reverberation time. The motivation to choose Eyring's formula is that it provides more accurate results in highly absorptive environments ($\bar{\alpha} > 0.3$) than the ones obtained with Sabine's formula, which assumes a low mean absorption coefficient in the room.
- Floor reflection, given by the term $Q^*/[4\pi(2d)^2]$ inside the logarithm, adding to the diffuse-field attenuation term. The floor reflection is considered present in all measurements. It is assumed that the floor is totally reflective and that the mouth and the ears are at a height of 1.5 m above the floor. All the early reflections from the walls, when averaged across positions in a room, are included in the diffuse-field attenuation term. The reflection from the ceiling is included in the diffuse-field attenuation term because the distance to the ceiling varies across rooms, and the model is valid on average in an ensemble of rooms.
- Diffuse-field HRTF (ΔL_{HRTF}), accounting for the increase in level associated to the use of a dummy head instead of a small omnidirectional microphone for the measurement of sound reflections.
- Direct sound characterization with the term $-K$, which represents the difference between the source sound power level and the SPL at the eardrum, empirically determined from laboratory measurements of these quantities using a head and torso simulator B& K type 4128 [18].

Table 3: Relevant frequency-dependent quantities used in the prediction model of voice support ST_V and decay time $DT_{40,ME}$. Reproduced from [18].

Center Frequency [Hz]	125	250	500	1000	2000	4000
K [dB]	3.6	3.7	2.5	4.0	8.3	8.1
Q^*	0.95	0.78	0.79	0.60	0.21	0.25
ΔL_{HRTF} [dB]	0	0	2	4	11	13
$L_{D,ref}$ [dB]	58.0	69.1	73.5	71.7	69.0	63.0

The frequency-dependent values of Q^* , ΔL_{HRTF} and K are given in Table 3. An overall value of voice support is obtained by weighting the frequency-dependent values with the typical speech spectrum at the ears of a speaker ($L_{D,ref}$ in Table 3):

$$ST_V = L_R - L_D = 10 \log \frac{\sum_{i=1}^6 10^{\frac{L_{D,ref,i} + ST_{V,i}}{10}}}{\sum_{i=1}^6 10^{\frac{L_{D,ref,i}}{10}}} \text{ (dB)}, \quad (4)$$

where the subindex i indicates an octave band value, among the octave bands with center frequencies from 125 Hz ($i = 1$) to 4 kHz ($i = 6$). The predictions from this model are shown in Fig. 4, considering a flat T_{60} across frequency and a room of proportions 2.8:1.6:1. ST_V decreases almost linearly with the logarithm of V (except for the largest volumes at low reverberation times) and increases with T_{60} . The axis on the right edge shows the average voice power level variations experienced by speakers, according to Eq. (2). The prediction model for ST_V has been validated with measurements in 30 educational facilities [18].

3.2 Decay time $DT_{40,ME}$

The decay time $DT_{40,ME}$ [1] (in which the subindex ME stands for Mouth-to-Ears) is defined as the time it would take for the backwards integrated energy curve of an OBRIR

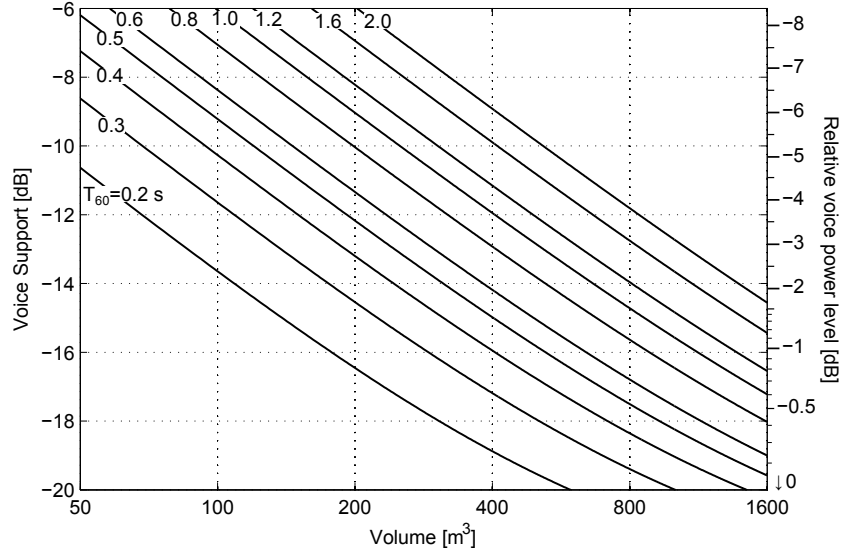


Figure 4: Average voice support versus room volume for different values of T_{60} , considering a flat T_{60} across frequency. The axis on the right edge shows the average voice power level variations experienced by speakers according to Eq. (2).

to decay 60 dB after the arrival of the direct sound, calculated from the initial decay of 40 dB (value of the subindex in $DT_{40,ME}$) and assuming a linear decay. Differently from traditional impulse responses in which the receiver is far away from the source, in an OBRIR source and receiver are located very close to each other, so the direct sound has much more energy than the reflected sound. Therefore, the $DT_{40,ME}$ is very sensitive to the direct-to-reflected sound level difference.

3.2.1 Relation with voice

A previous study analyzed relevant subjective attributes related to the experience of talking in different rooms using questionnaires [1]. It was found that $DT_{40,ME}$ is linearly related to the sensation of reverberance. More interestingly, the general sensation of *vocal comfort* \hat{C} is non-linearly related to $DT_{40,ME}$ for all speakers with normal voices. The vocal comfort \hat{C} is a weighted average of the answers to four different questions

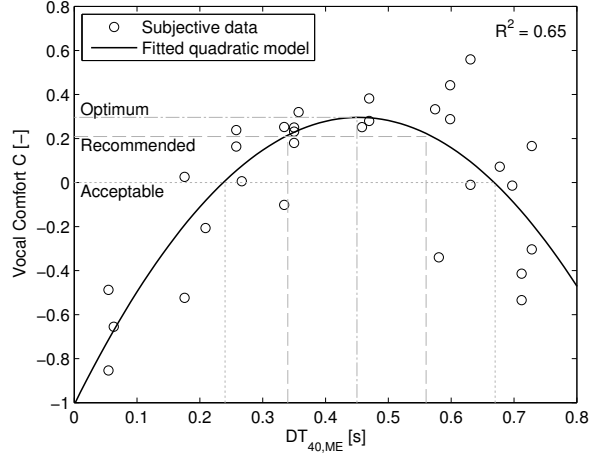


Figure 5: Vocal comfort \hat{C} as a function of $DT_{40,ME}$ in the classrooms where subjects were asked to talk during an experiment. A quadratic regression model is shown, defining an optimum point and ranges of recommended and acceptable $DT_{40,ME}$ values.

regarding 1) the general evaluation of the room for speech purposes, 2) the perceived support, 3) the feeling of having to raise the voice and 4) the sensation of tiredness after speaking longly in the room. The units of \hat{C} are standard deviations across scores for the same individual, whereas the global average vocal comfort is expressed as $\hat{C} = 0$. The pooled results of \hat{C} for different acoustic conditions represented by $DT_{40,ME}$ are shown in Fig. 5. Each point in the figure represents the average \hat{C} given by a group of participants for one condition in the experiment. There were three groups of participants, each containing between 11 and 13 participants.

A quadratic regression model ($R^2 = 0.65; p < 0.001$) of the pooled results of \hat{C} for speakers with normal voices is shown in Fig. 5 and defined by

$$\hat{C} = -6.4 DT_{40,ME}^2 + 5.8 DT_{40,ME} - 1.0, \quad (5)$$

The optimum vocal comfort \hat{C}_{opt} is located at $DT_{40,ME,opt} \approx 0.45$ s. Three categories of vocal comfort are defined: recommended, acceptable, and non-acceptable. The recommended vocal comfort range is where $\hat{C} \geq \hat{C}_{opt}/\sqrt{2}$, thus for $DT_{40,ME}$ between 0.35 and

0.55 s. In the acceptable range, $0 \leq \hat{C} \leq \hat{C}_{\text{opt}}/\sqrt{2}$, which occurs at two different intervals: $\text{DT}_{40,\text{ME}}$ from 0.25 to 0.35 s and from 0.55 s to 0.65 s. The non-acceptable range is given by a less than average vocal comfort ($\hat{C} \leq 0$), which occurs for $\text{DT}_{40,\text{ME}} < 0.25$ s and for $\text{DT}_{40,\text{ME}} > 0.65$ s.

3.2.2 Prediction model

The prediction model for the average $\text{DT}_{40,\text{ME}}$ in a room [1] makes the same assumptions and contains the same elements as the model for ST_V but taking into account the time structure of the OBRIR. The basic model of an OBRIR, considered for the $\text{DT}_{40,\text{ME}}$ prediction model, is shown in Figure 6 and consists of three components:

- Direct sound, modeled as a Dirac delta at $t = 0$, $\delta(t)$
- Floor reflection, modeled as a delayed and attenuated Dirac delta, $A\delta(t - t_0)$, with parameters

$$10 \log A = 10 \log \left(\frac{Q^*}{4\pi(2d)^2} \right) + \Delta L_{\text{HRTF}} - K \text{ (dB)}, \quad (6)$$

$$t_0 = \frac{2d}{c} \text{ (s)}, \quad (7)$$

- A reverberant tail, modeled as a decaying exponential function, $Be^{-t/\tau}$, with parameters

$$10 \log B = 10 \log \left(\frac{4(1 - \bar{\alpha})}{S\bar{\alpha}} \right) + \Delta L_{\text{HRTF}} - K - 10 \log \tau \text{ (dB)}, \quad (8)$$

$$\tau = \frac{T_{60} \log e}{3} \text{ (s)}, \quad (9)$$

In these equations, $\bar{\alpha}$ is the mean absorption coefficient of the room obtained through Eyring's formula, S the total surface area, Q^* the directivity of a speaker in the downward direction, d the distance from the mouth to the floor ($= 1.5$ m), ΔL_{HRTF} the magnitude of the diffuse-field head-related transfer function (HRTF, in dB), and K the

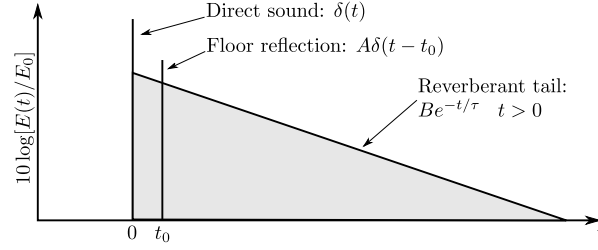


Figure 6: Energy density time curve of the OBRIR assumed in the prediction model for $DT_{40,ME}$, showing the main components in the airborne acoustic path between the mouth and the ears: the direct sound, the floor reflection, and the reverberation tail

difference between the SPL at the eardrum and L_W . The values of Q^* , ΔL_{HRTF} , and K are shown in Table 3.

One of the differences with the prediction model for ST_V is that the model for $DT_{40,ME}$ does not have a closed mathematical expression and is computed by means of an algorithm with the following steps: modelling of a parametric OBRIR (Fig. 6), calculation of the backward integrated energy curve, search of the time instant t_{-40} when the backwards integrated energy curve decays -40 dB relative to $t = 0$, and finally $DT_{40,ME} = 1.5 t_{-40}$.

In order to assess the prediction model, a retrospective analysis of the measurements in 30 educational facilities [18] was performed. The details of the validation are shown in the appendix. According to this, the model yields the most accurate predictions for the octave bands of 2 kHz and 4 kHz. The predictions of the model for the average in the octave bands of 2 kHz and 4 kHz, considering a flat T_{60} across frequency and a room of proportions 2.8:1.6:1, are shown in Fig. 7, as a function of T_{60} and V . For typical room volumes, $DT_{40,ME}$ is larger than T_{60} and decreases with V .

An approximate regression model fitted to the predictions of the above model in the range of volumes between 50 and 1600 m³ and reverberation times between 0 and 1.6

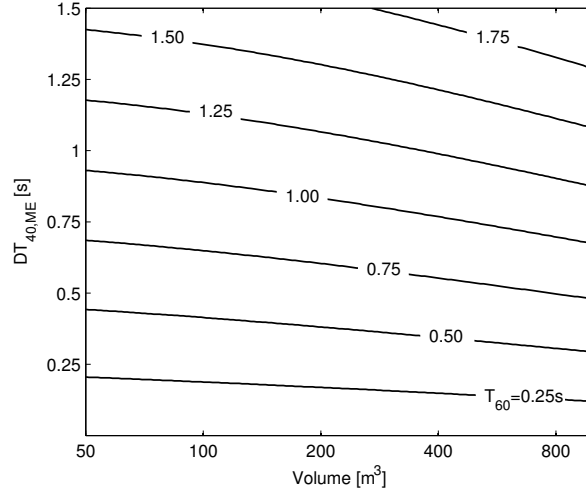


Figure 7: Mean decay time $DT_{40,ME}$ versus volume for different values of T_{60} , obtained as the average of the octave band values of 2 kHz and 4 kHz.

s, which is more convenient for calculations, is given by the expression

$$DT_{40,ME} \approx 1.33 T_{60} - 0.022 \log_{10} V - 0.19 T_{60} \times \log_{10} V \text{ (s)}. \quad (10)$$

4 Speaker-oriented classroom acoustics guidelines

The previous section described the room acoustic parameters relevant for a speaker and their relationship with voice. The present section aims at applying those relationships to the practical design of classroom acoustics.

4.1 Optimization of vocal comfort in classrooms

From the non-linear relationship between vocal comfort and $DT_{40,ME}$, shown in Eq. (5) and in Fig. 5, the range of $DT_{40,ME}$ between 0.35 s and 0.55 s was found to provide recommended vocal comfort levels, whereas the extended range from 0.25 s to 0.65 s provided acceptable vocal comfort. These range of values are shown in Fig. 8, which is a replot of the prediction model of $DT_{40,ME}$ of Fig. 7, represented as a function of

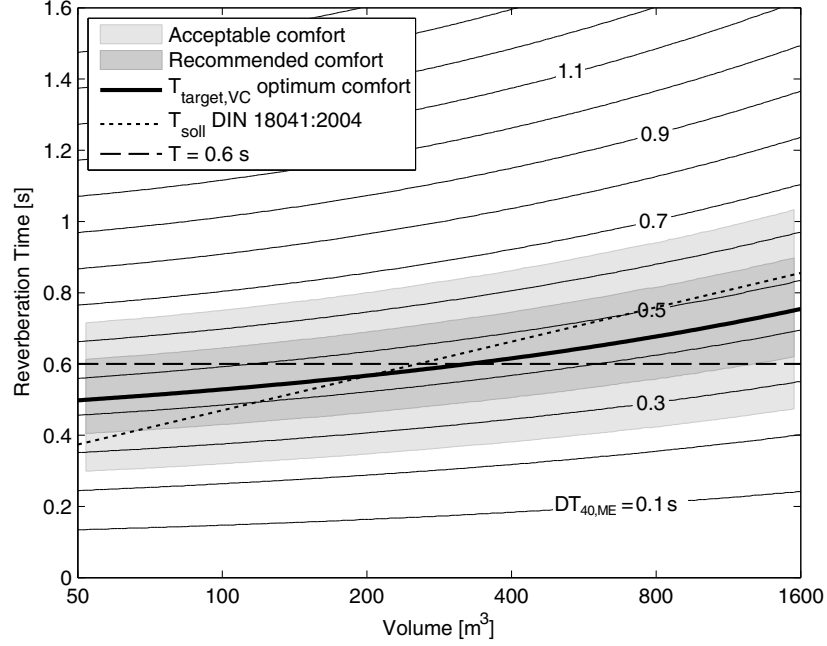


Figure 8: Level sets of the predicted $DT_{40,ME}$ (units are s), as a function of reverberation time and volume. The areas of $DT_{40,ME}$ providing recommended and acceptable voice comfort levels are indicated in darker and lighter gray color. Bold thick line: target reverberation time $T_{\text{target,VC}}$ which maximizes the vocal comfort. Dotted line: target reverberation time T_{soll} in classrooms defined in the German standard DIN18041:2004. Dashed line: volume-independent reverberation time of 0.6 s.

reverberation time and volume. In this way, it is straightforward to compare the areas of recommended and acceptable vocal comfort with the design recommendations of standards like the German DIN 18041:2004 [20], which defines a target reverberation time T_{soll} as a function of volume (see Table 2) or with other recommendations that set a volume-independent maximum reverberation time.

Figure 8 shows also the target reverberation time which optimizes the vocal comfort. A suitable empirical fit for the target reverberation time is given by

$$T_{\text{target,VC}} = 0.032\sqrt[3]{V} + 0.38 \text{ (s)}, \quad (11)$$

In principle, the reverberation time in classrooms for optimum vocal comfort should be between 80% and 120% of $T_{\text{target,VC}}$. This variation range falls within the recommended vocal comfort range, i.e. in the range $0.8T_{\text{target,VC}} \leq T \leq 1.2T_{\text{target,VC}}$, $0.35 \text{ s} \leq \text{DT}_{40,\text{ME}} \leq 0.55 \text{ s}$.

In European standards, one trend is set by the German DIN 18041:2004, in which the target reverberation time T_{soll} depends linearly on the logarithm of the volume (with an offset). Other standards define a volume-independent reverberation time requirement. In Fig. 8, T_{soll} and a constant T of 0.6 s are shown for comparison purposes, together with $T_{\text{target,VC}}$. The reverberation time for optimum vocal comfort $T_{\text{target,VC}}$ depends on the volume of the room, but does not follow a straight line in a volume-logarithmic scale, and has a slope which is approximately half of the slope of T_{soll} . For volumes lower than 200 m^3 , $T_{\text{target,VC}}$ is larger than T_{soll} . However, for volumes larger than 200 m^3 , $T_{\text{target,VC}}$ is lower than T_{soll} . Nevertheless, the differences between $T_{\text{target,VC}}$ and T_{soll} are under 0.1 s. Similar differences are obtained by comparing $T_{\text{target,VC}}$ with the constant $T = 0.6 \text{ s}$, but in this case $T_{\text{target,VC}}$ is below 0.6 s for volumes lower than 300 m^3 and above 0.6 s for larger volumes. In terms of vocal comfort, the German DIN 18041:2004 does not define a more convenient target reverberation time than a volume-independent T of 0.6 s in occupied conditions, but its main advantage over other standards is the prescription of target values rather than maximum limit values.

4.2 Assessment of vocal effort

The recommendations presented above have been derived from a vocal comfort maximization criterion. Thus, the vocal effort has not been introduced in these guidelines. Figure 9 shows the mutual relationship between ST_V and $\text{DT}_{40,\text{ME}}$, for equal values of V (dotted lines) and T_{60} (solid lines). This chart is obtained by merging the prediction models shown in Figs. 4 and 7, as both models have been described uniquely in terms of V and T_{60} . This is meaningful because the horizontal axis (ST_V) is linked to the

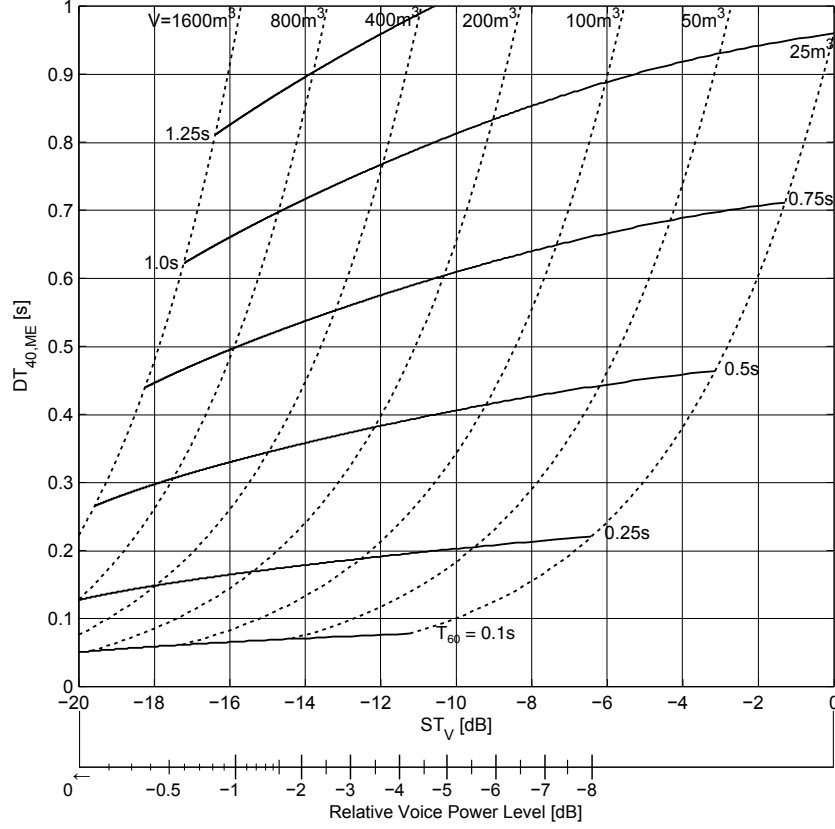


Figure 9: Decay time $DT_{40,ME}$ versus voice support ST_V for different values of T_{60} (solid lines) and V (dotted lines). The bottom axis shows the average voice power level variations experienced by speakers, according to Eq. (2).

vocal effort, whereas the vertical axis ($DT_{40,ME}$) is related to the *vocal comfort*. On the bottom axis of the figure, there is an indication of ΔL_W experienced by a speaker in the presence of low noise levels ($\Delta L_W = 0$ dB corresponds to the average L_W produced in an anechoic room) according to Eq. (2). The values in this axis illustrate how different classroom acoustic designs affect the voice levels of teachers while the audience is silent.

To illustrate the use of the chart, an example is given. We would like to assess the effect of refurbishing a classroom of 200 m^3 in terms of vocal effort and vocal comfort. Initially, the reverberation in the classroom is 1.0 s . For this condition, we would

look in Fig. 9 for the intersection for the level lines of $V = 200 \text{ m}^3$ and $T_{60} = 1.0$ s, and find that $DT_{40,ME,0} \approx 0.83$ s, $ST_{V,0} \approx -9.0$ dB, and $\Delta L_{W,0} \approx -6.0$ dB. One possible refurbishment would lower the reverberation time to 0.75 s. In that case, $DT_{40,ME,1} \approx 0.6$ s, $ST_{V,1} \approx -10.5$ dB, and $\Delta L_{W,1} \approx -5.0$ dB. Here, the refurbishment would enhance the vocal comfort from “non-acceptable” to “acceptable”. The vocal effort would increase, as teachers would raise their voices by approximately $\Delta L_{W,1} - \Delta L_{W,0} \approx 1.0$ dB. Another possible refurbishment would lower the reverberation time to 0.5 s in order to fulfill usual design requirements. In that case, $DT_{40,ME,2} \approx 0.38$ s, $ST_{V,2} \approx -12.3$ dB, and $\Delta L_{W,2} \approx -3.5$ dB. In this case, the vocal comfort would be still acceptable but the vocal effort would be increased with respect to the initial situation by approximately $\Delta L_{W,2} - \Delta L_{W,0} \approx 2.5$ dB.

It is apparent from Fig. 9 that lower reverberation times lead to an increased vocal effort. However, there are no studies which link these variations in vocal effort to an increased prevalence of vocal problems among teachers. If this data were available, it would be possible to refine the classroom acoustics guidelines with vocal effort considerations, setting a limit to the minimum recommended ST_V . For example, if a study found that an acceptable value of ΔL_W was more than -3 dB, then a good classroom acoustic design should set ST_V higher than -13 dB.

5 Assessment of guidelines using speech intelligibility metrics

Any recommendation for classroom acoustic design should take into account students’ needs in terms of speech intelligibility. For this, the recommendations of the previous section will be assessed with a model of speech intelligibility according to the early-to-late ratio C_{50} and a more advanced model which uses the useful-to-detrimental ratio U_{50} , based on an empirical model of actual signal-to-noise ratios (SNR) in classrooms.

5.1 Basic prediction model of speech intelligibility using C_{50}

The early-to-late ratio C_{50} , sometimes referred to as speech clarity, can be used as a rough indicator of speech intelligibility, empirically related to the speech transmission index (STI) through the basic relationship [41]

$$\text{STI} = 0.030C_{50} + 0.555. \quad (12)$$

The C_{50} is defined as the level difference of the early arriving speech $L_{\text{SL,early}}$ and the late arriving speech $L_{\text{SL,late}}$ signals. Early arriving speech enhances speech intelligibility, whereas late arrivals are detrimental, and therefore high values of C_{50} are in principle beneficial. Thus,

$$C_{50} = L_{\text{SL,early}} - L_{\text{SL,late}} \text{ (dB)}. \quad (13)$$

According to Nijs and Rychtarikova [41], $L_{\text{SL,early}}$ contains the direct sound and the early reflections arriving in the 50 ms after the arrival of the direct sound and is given by

$$L_{\text{SL,early}} = L_W + 10 \log \left(\frac{Q}{4\pi r^2} + \frac{4(1 - \bar{\alpha})^{\beta_{\text{fb}} \cdot r / l_{\text{mfp}}}}{\bar{\alpha} S} (1 - e^{-0.69/T_{60}}) \right) \text{ (dB)}, \quad (14)$$

where Q is the directivity of the speaker in the direction aiming at the listener, r is the distance between speaker and listener, $\bar{\alpha}$ is the mean absorption coefficient of the room, β_{fb} is an experimental parameter related to the spatial decay rate of the reverberant field and l_{mfp} is the mean free path of the room ($l_{\text{mfp}} = 4V/S$). It is assumed that the reflected sound follows an exponentially decaying random process with an amplitude that decays with longer distances to the source, as suggested by Barron in the context of acoustics for performance spaces (without any parameter β_{fb} , i.e. $\beta_{\text{fb}} = 1$) [42] and experimentally refined by Sato and Bradley [43] in the context of classroom acoustics by introducing the parameter β_{fb} , usually > 1 . The motivation to use this parameter $\beta_{\text{fb}} > 1$ was to account for a faster decay of SPL with distance beyond the reverberation distance in spaces with higher scattering (e.g. classrooms) than in performance spaces.

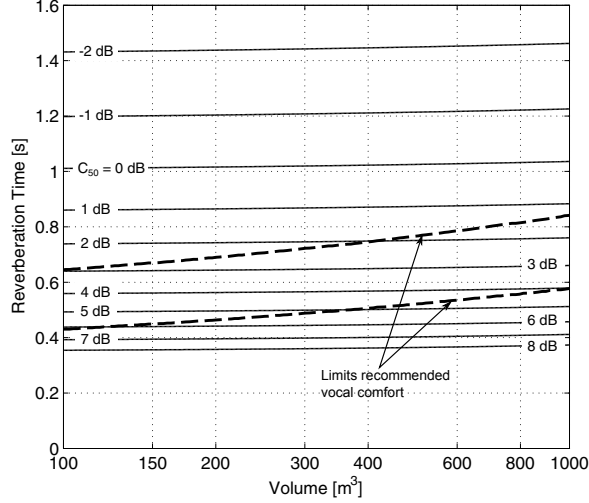


Figure 10: C_{50} as a function of the room volume and reverberation time, considering the model in Eqs. (13)–(15). The upper and lower limits of reverberation time to achieve recommended vocal comfort levels are indicated in bold dashed lines.

The late arriving speech level $L_{\text{SL,late}}$, on the other hand, contains the effect of sound reflections arriving later than 50 ms after the arrival of the direct sound:

$$L_{\text{SL,late}} = L_W + 10 \log \left(\frac{4(1 - \bar{\alpha})^{\beta_{\text{fb}} \cdot r / l_{\text{mfp}}}}{\bar{\alpha} S} e^{-0.69/T_{60}} \right) \text{ (dB)}. \quad (15)$$

Figure 10 shows the C_{50} as a function of room volume and reverberation time. It is assumed that the rooms are rectangular with dimension ratios 2.8:1.6:1 (length : width : height) and equal reverberation time across frequency. For simplicity, it is assumed that the speaker can be oriented toward any direction (thus, $Q = 1$), that the speaker-to-listener distance is the geometrical mean of the length and the width of the room, and that $\beta_{\text{fb}} = 1$, as in [41]. Moreover, $\bar{\alpha}$ is calculated statistically from Eyring’s formula. It can be observed that C_{50} varies weakly with volume and depends mostly on reverberation time. It is therefore that classroom acoustics standards setting a volume-independent maximum limit of reverberation time seek to maximize speech intelligibility through the C_{50} metric.

In Fig. 10, the upper and lower limits of reverberation time that result in recom-

mended vocal comfort levels are shown with bold dashed lines. This range of reverberation time provides speech clarity higher than 1.5 dB (or $STI > 0.60$), which is sometimes used as a design requirement for good speech intelligibility (e.g. [31]). A value of reverberation time around 0.6 s for all typical classroom volumes is within the recommended values of reverberation time for vocal comfort (see Fig. 8) and therefore provides also a good speech intelligibility. It should also be taken into account that among children, hearing impaired persons and students of a foreign language, speech intelligibility is lower than for normal listeners, and thus the lowest reverberation time given by the bold dashed line ($0.8T_{\text{target,VC}}$) should be preferred over higher values.

5.2 Signal-to-noise ratios in classrooms

Speech intelligibility is degraded due to noise, which is commonly present in educational settings. Noise is usually quantified relative to the strength of the signal, i.e. the signal-to-noise ratio (SNR). A more formal definition of SNR is the difference between the sound power levels of speech $L_{W,\text{speech}}$ and student-activity noise $L_{W,\text{SA}}$,

$$\text{SNR} = L_{W,\text{speech}} - L_{W,\text{SA}} \text{ (dB)}. \quad (16)$$

Empirical models that link student-activity noise, instructor voice power levels, geometric and acoustic data of the room and the number of students in university classrooms are given by Hodgson *et al* [17]. The inputs to the models were measurements during 18 lectures in 11 classrooms with volumes between 110 and 957 m³, between 6 and 254 students present in the classroom, and total occupied absorption areas between 30 and 305 m². The A-weighted student-activity noise level L_{SA} is described as:

$$L_{\text{SA}} = 83 + 10 \log N - 34.4 \log A_0 + 0.08 A_0 \text{ (dB re } 20\mu\text{Pa)} \quad (17)$$

where A_0 is the total absorption area in the room ($A_0 = S\bar{\alpha}$). The A-weighted student-activity noise power level $L_{W,\text{SA}}$ is

$$L_{W,\text{SA}} = L_{\text{SA}} - 10 \log \left(\frac{4(1 - \bar{\alpha})}{S\bar{\alpha}} \right) \text{ (dB re 1 pW)} \quad (18)$$

where it is assumed that L_{SA} is homogeneous throughout the classroom.

The A-weighted sound power levels of speech $L_{W,\text{speech}}$, averaged for male and female teachers, is [17]:

$$L_{W,\text{speech}} = 53.5 + 0.5 L_{SA} + 0.016 V - 9.6 \log A_0 \text{ (dB re 1 pW)}. \quad (19)$$

Due to the Lombard effect, $L_{W,\text{speech}}$ increases with L_{SA} . $L_{W,\text{speech}}$ depends also on the volume V and the total absorption area A_0 . As an empirical model in actual noise conditions, Eq. (19) is not related to the model described in Eq. (2), which describes variations of L_W in quiet conditions.

5.3 A prediction model of speech intelligibility using a useful-to-detrimental ratio

The useful-to-detrimental ratio U_{50} is a more suitable measure of speech intelligibility than C_{50} because it takes into account the presence of noise. In this case, the useful part is the sound signal arriving in the first 50 ms after the arrival of the direct sound (including it) and the detrimental part contains the late reflections and the noise. Thus,

$$U_{50} = L_{SL,\text{early}} - L_{SL,\text{late+noise}} \text{ (dB)}. \quad (20)$$

The early arriving speech level $L_{SL,\text{early}}$ is given in Eq. (14), whereas $L_{SL,\text{late+noise}}$ according to Nijs and Rychtarikova [41] is

$$L_{SL,\text{late+noise}} = L_W + 10 \log \left(\frac{4(1 - \bar{\alpha})^{\beta_{fb} \cdot r / l_{mfp}}}{\bar{\alpha} S} e^{-0.69/T_{60}} + \frac{4 \cdot 10^{-\text{SNR}/10}}{\bar{\alpha} S} \right) \text{ (dB)}, \quad (21)$$

which differs from Eq. (15) in the term containing the SNR.

Using the SNR description from Eqs. (16)–(19), which models typical SNR in classrooms as a function of the acoustic conditions, the resulting U_{50} depends on the number of students N present in the classroom. Figure 11 shows the U_{50} values as a function of the volume, the reverberation time and the number of students ($N = 10$,

20, 40, and 80). The horizontal axis has been adjusted to display volumes of at least 4 m^3 per student. Lower volumes per student might be problematic in terms of indoor air quality, as the flow of air to be exchanged would increase, and might make teachers and students *feel overwhelmed, discouraged and often disgusted* due to overcrowding [44]. Identically to the model for C_{50} , it is assumed that the rooms are rectangular with dimension ratios 2.8:1.6:1 (length : width : height) and equal reverberation time across frequency. For simplicity, it is assumed that the speaker can be oriented toward any direction ($Q = 1$), that the speaker-to-listener distance is the geometrical mean of the length and the width of the room, and that $\beta_{\text{fb}} = 1$. In the figure, the equal U_{50} contours are shown in steps of 1 dB. In the same figure, the upper and lower limits of reverberation time that result in recommended vocal comfort levels are shown with bold dashed lines. The hashed areas in Fig. 11 indicate areas for which the prediction of SNR is not valid due to unavailable data for the model. The values of U_{50} are substantially lower than those of C_{50} in the same conditions of volume and reverberation time, and decrease with increasing number of students, due to poorer SNRs. Differently from Fig. 10, the U_{50} values worsen for reverberation times below the ones given by the maximum- U_{50} curves in Fig. 11. The reverberation time curve that defines maximum U_{50} as a function of the volume is close to the lower limit of recommended vocal comfort zone ($0.8T_{\text{target,VC}}$) for 10 or 20 students. This curve increases with the number of students and falls within the recommended range of reverberation time that optimizes vocal comfort.

None of the U_{50} values in Fig. 11 exceeds 3 dB (i.e. $\text{STI} > 0.65$), even with a low number of students. Nevertheless, the student-activity noise levels used for the SNR calculations are median levels of the periods when the teacher is not talking. If effective noise control measures are taken, so that there are at least 15 dB between background noise (without activity) and instructor speech level, actual speech intelligibility can be expected to be somewhere in between the predictions of Figs. 10 and 11, as C_{50} would be a reasonable predictor of speech intelligibility in the best scenario when students are

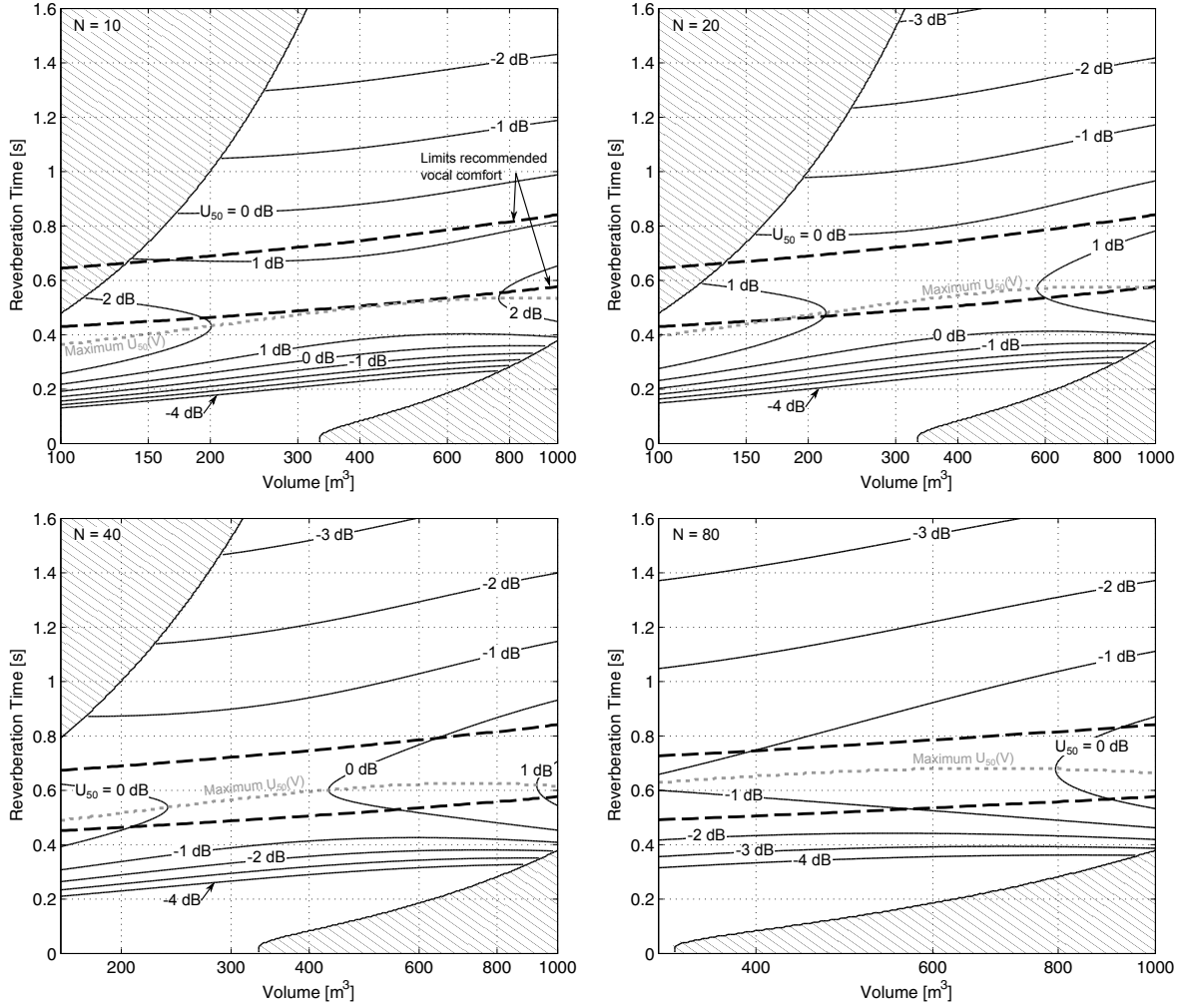


Figure 11: U_{50} as a function of room volume and reverberation time, considering typical SNR measured in classrooms, for $N = 10$ students (top left), $N = 20$ (top right), $N = 40$ (bottom left) and $N = 80$ (bottom right). The upper and lower limits of reverberation time to achieve recommended vocal comfort levels are indicated in bold dashed lines. Gray dotted lines indicate reverberation times that maximize U_{50} for each volume. Hatched areas correspond to combinations of volume and reverberation time for which the SNR predictions are not valid.

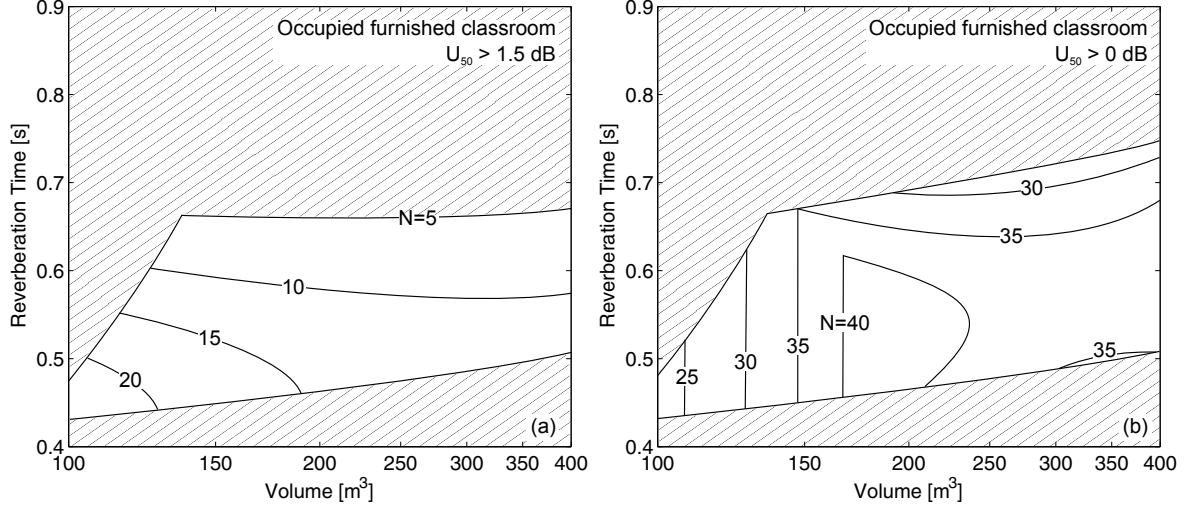


Figure 12: Areas of recommended volume and reverberation time in occupied classrooms, for a given number of students N , which satisfy simultaneously criteria of vocal comfort and speech intelligibility for (a) $U_{50,\min} = 1.5$ dB and (b) $U_{50,\min} = 0$ dB. For each N , the design area should be considered as the interior of the boundary line.

silent.

A classroom acoustic design that maximizes simultaneously criteria of vocal comfort and speech intelligibility is found as the intersection of areas of recommended vocal comfort ($0.8T_{\text{target,VC}} \leq T \leq 1.2T_{\text{target,VC}}$) and the areas that provide speech intelligibility higher than a minimum value ($U_{50} \geq U_{50,\min}$). If, for example, $U_{50,\min}$ is taken as 1.5 dB ($\text{STI} > 0.6$), the overlap between areas decreases with the number of students, and there is no overlap for 40 or 80 students. This overlap is summarized in Fig. 12(a) as target values of volume and target reverberation times in occupied classrooms as a function of the number of students. The area of recommended design decreases with the number of students in the classroom, and beyond $N = 20$ students, the speech intelligibility criteria becomes too restrictive. In this case, the best acoustic conditions for small groups of students (up to 20 pupils) are obtained for T values around 0.45 s (and volumes up to 120 m³) in occupied classrooms.

The choice of 1.5 dB for $U_{50,\min}$ leads to small groups of students, which is advisable for groups with special needs (e.g. students with hearing impairment or non-native speakers). A more realistic choice of $U_{50,\min} = 0$ dB ($\text{STI} > 0.55$) is used to derive Fig. 12(b), which shows the combinations of volume and reverberation time that provide simultaneously good vocal comfort and satisfactory speech intelligibility ($U_{50} > 0$ dB) for any number of students up to 40. For a classroom for flexible teaching with 40 students, design possibilities become restricted to reverberation times between 0.45 and 0.6 s and volumes between 160 and 210 m³. For less students, large volumes are not a restriction, besides the facts that it will be more expensive to achieve the target reverberation times (having to install more absorptive material) and that space is a scarce resource in many countries, especially in cities.

5.4 Guidelines for unoccupied furnished classrooms

Whereas the derivation of guidelines has been suggested for fully occupied conditions, regulations in classroom acoustics are most often referred to unoccupied but furnished conditions. In an occupied classroom with volume V , total surface area S and reverberation time T_{occ} , the mean absorption coefficient $\bar{\alpha}_{\text{occ}}$ according to Eyring's formula is

$$\bar{\alpha}_{\text{occ}} = 1 - \exp[-0.161V/(T_{\text{occ}}S)]. \quad (22)$$

In order to obtain an estimate of the acoustic conditions in unoccupied furnished classrooms, the absorption corresponding to the students and the teacher, considering that each person in the classroom absorbs $A_s = 0.28$ m² (according to Sato and Bradley [43]), is subtracted from the total absorption area in the occupied classroom $A_{\text{occ}} = S\bar{\alpha}_{\text{occ}}$. Thus, the mean unoccupied absorption coefficient $\bar{\alpha}_{\text{unocc}}$ results in

$$\bar{\alpha}_{\text{unocc}} = \frac{A_{\text{occ}} - (N + 1)A_s}{S}, \quad (23)$$

which leads to a reverberation time in unoccupied but furnished conditions T_{unocc} of

$$T_{\text{unocc}} = \frac{0.161V}{-S \ln(1 - \bar{\alpha}_{\text{unocc}})}. \quad (24)$$

Each of the curves limiting an area of recommended design in Fig. 12 is transformed into a curve which expresses the recommended design areas of volume and reverberation time in unoccupied furnished classrooms for a fixed number of students by applying Eqs. (22)–(24). It should be noted that there are areas which do not overlap for different number of students, as a classroom with high occupancy might in principle be allowed to have higher unoccupied T_{unocc} values, but which would result in excessive occupied T_{occ} values for low occupancy. For this reason, Fig. 13 shows the areas of recommended design in unoccupied but furnished classroom for a variable number of students, i.e. the areas that fulfill vocal comfort and speech intelligibility criteria simultaneously for any number of students between a minimum of $N_{\text{min}} = 5$ and a maximum N .

According to Fig. 13(a), reverberation times in unoccupied furnished classrooms designed for high speech intelligibility should not exceed 0.7 s in any case, whereas the minimum recommended reverberation time depends on the number of students. Such a classroom with up to 20 students should have a T_{unocc} of around 0.55 s. In classrooms with less strict speech intelligibility requirements, as shown in Fig. 13(b), the maximum T_{unocc} varies almost linearly with the logarithm of the volume from 0.7 s at 120 m³ to 0.75 s at 400 m³. The minimum T_{unocc} for classrooms with $N \leq 20$ students is slightly higher than 0.5 s, whereas the minimum T_{unocc} for classrooms with $N \leq 40$ students is 0.6 s.

6 Discussion

A classroom acoustic design based only on vocal comfort suggests that the target reverberation time $T_{\text{target,VC}}$ in conditions of full occupancy is dependent on the cubic root of the volume of the classroom. As shown in Fig. 8, these target values—when

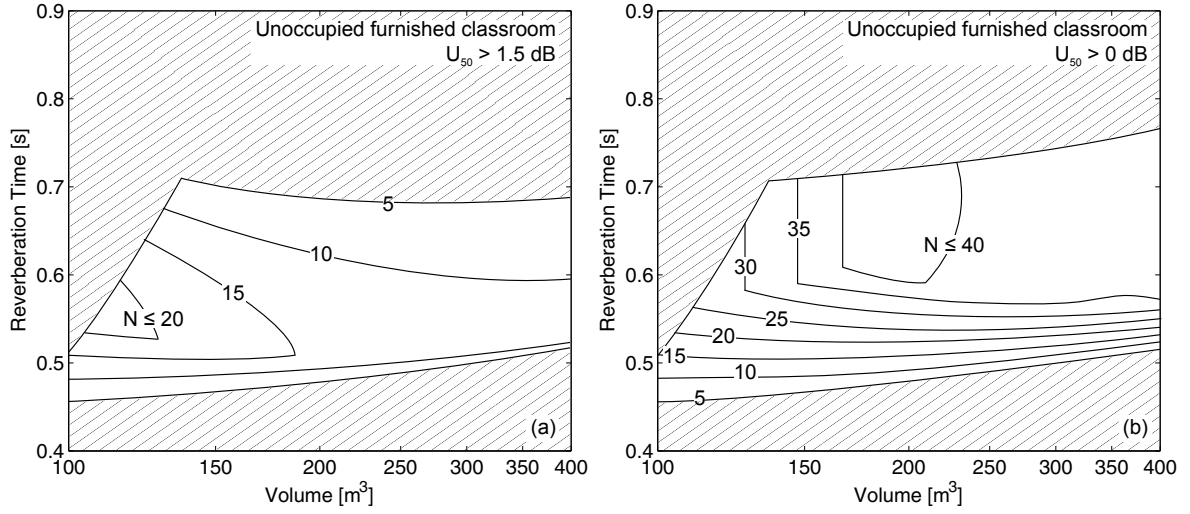


Figure 13: Areas of recommended volume and reverberation time in unoccupied but furnished classrooms, for a given number of students (between 5 and N), which satisfy simultaneously criteria of vocal comfort and speech intelligibility for (a) $U_{50,\min} = 1.5$ dB and (b) $U_{50,\min} = 0$ dB. For each N , the design area should be considered as the interior of the boundary line.

represented against the volume in a logarithmic scale—have a slope that is in between that of the target values in the German DIN18041:2004 and that of a flat target T . Moreover, Fig. 11 showed that the T that maximized the useful-to-detrimental ratio U_{50} , which takes into account the signal-to-noise ratio and the early and late arriving reflections in a classroom, had a similar slope as $T_{\text{target,VC}}$. This is an argument in favor of defining volume-dependent target reverberation times in future revisions of regulations and recommendations on classroom acoustics but with a lower slope than that defined on the DIN18041:2004.

Classrooms are spaces for listening as well as for speaking, and therefore speech intelligibility criteria has also been taken into account, using U_{50} as the metric, to derive a more refined set of recommendations. The signal-to-noise ratios used as an input for the U_{50} prediction model are based on measurements of activity noise and voice levels by Hodgson *et al* [17] in university classrooms. It is likely that levels measured at secondary and primary school classrooms might differ from these, because older students tend to be more silent than younger ones, according to one study by Picard and Bradley [45]. However, other studies do not find such a relationship, as pointed out by Shield and Dockrell [46], except for day-care centers and kindergarden, where activity noise levels are significantly higher. To the knowledge of the authors at the current moment, there are no available empirical models which link speech levels, activity noise levels, number of students and room acoustic conditions for younger pupils, and thus the data by Hodgson *et al* provides very valuable information. For this reason, the models in Eq. (17) and Eq. (18) might suffer some variations that could lead to slightly different design guidelines based on U_{50} . Moreover, the contribution of different frequency bands to the sensation of speaking support or vocal comfort has not been studied. There could be some dominant bands, as it is the case of e.g. speech intelligibility, where the octave bands of 500 Hz, 1 kHz and 2 kHz are the most important ones [47]. Another important detail is that the optimum vocal comfort was determined in laboratory with

very low noise levels, which may be different from the actual situation in primary and secondary school classrooms. The dependence of activity noise with room acoustic conditions or room occupancy was not characterized in the laboratory experiments, but the latter can be representative of quiet conditions in classrooms under a climate of discipline among students (with student-activity noise levels below 45 dB, which is the level below which the teacher does not adjust the vocal intensity according to the Lombard effect [40]). It is hypothesized that, if the teacher is not forced to modify vocal intensity, preferred $DT_{40,ME}$ values will remain equal). Despite these shortcomings, the recommendations for classroom acoustics design derived in the previous section and summarized in Figs. 12 and 13 constitute one of the first attempts to combine speaker and listener requirements. They should encourage better classroom acoustic design and further research that improves the current understanding on this topic.

These recommendations represent average values in classrooms and are meant for situations where the position of the speaker is variable, as in flexible teaching methods. In this case, no acceptable acoustic conditions can be achieved for more than approximately 40 students without exposing the teacher to talk uncomfortably or students to experience noticeably degraded speech intelligibility. The recommended values of reverberation time in classrooms with up to 40 students are between 0.45 s and 0.6 s and the volumes should be between 160 and 210 m³. In unoccupied conditions, the same classrooms should have reverberation times between 0.6 and 0.7 s. For less students, paradoxically, higher volumes are possible within the recommendations, as shown in Figs. 12 and 13. The reason for this is that signal-to-noise ratio decreases with increasing number of students, and design possibilities become more restricted.

It should be noted that younger children, hearing impaired persons and students of a foreign language achieve lower speech intelligibility scores than normal adult listeners under the same acoustic conditions and thus the design criteria of Figs. 12(a) and 13(a) (more restrictive speech intelligibility criteria, $U_{50,min} = 1.5$ dB) should be preferred

over Figs. 12(b) and 13(b) (with $U_{50,\min} = 0$ dB). In this case, there are more severe restrictions on the maximum number of students in the classroom (which should not exceed 20) and the maximum volume. In addition, the lowest reverberation times values should be preferred among the different design options.

The limitations in volume, number of students and reverberation time on the above optimum classroom acoustic conditions is valid only for flexible teaching methods. For larger number of students, teaching will probably take place in the form of lectures, with defined positions for teacher and students. In this case, a specific design should be made taking into account the position of the teacher. Only by considering the directivity of the speaker in the direction of the students ($Q > 1$), U_{50} values in Figs. 10 and 11 will increase and show a larger overlap with the recommended vocal comfort. If, additionally, the surfaces around the teacher are designed to enhance early reflections, these will also contribute to higher values of C_{50} and U_{50} for the same conditions of volume and reverberation time in classrooms. Moreover, ST_V will also increase due to the enhanced early reflections (which increase the reflected sound level L_R without modifying the direct sound level L_D) and the vocal effort will be lowered. In making such a design, the values of $DT_{40,ME}$ and ST_V will be highly dependent on the position and orientation of the speaker in the room and should be evaluated with more advanced prediction tools (e.g. room acoustics simulation software packages based on geometrical acoustics). Based on these observations, when designing classrooms larger than 200 m^3 and for more than 40 students, a dedicated acoustic study should be made, in which the target reverberation time in the occupied classroom should be $T_{\text{target,VC}}$, as expressed in Eq. (11), which depends on the cubic root of the volume.

The presence of long reverberation times and too high noise levels increases the annoyance among teachers. This effect has been studied by Kristiansen *et al.* [4], using self-reported estimates of noise exposure and disturbance attributed to noise among 419 secondary school teachers in 10 schools in Copenhagen. The reverberation time

was measured in the classrooms, and the schools were grouped in three categories: short (0.38 - 0.46 s), medium (0.46 - 0.66 s) and long (0.61 - 0.86 s) reverberation time. The reverberation time categories played a significant role for the disturbance attributed to noise from children in the class, and there was also a tendency that long reverberation times were associated with higher self-reported noise exposure. The study also found that the degree of noise attributed to children in the class had a stronger impact than sole reverberation time on self-reported noise exposure. Moreover, the number of children in the classroom was significantly associated with self-reported noise exposure. Thus, a way of having a good environment for both teachers and students, with low levels of noise, is by keeping low the number of students in each classroom.

The recent Essex study [48] observed that A-weighted activity noise levels in classrooms decrease 9 dB per 0.4 s of reduction in reverberation time, whereas teachers' speech levels decreased at a much lower rate, 4 dB per 0.4 s of reduction in reverberation time. Thus, in those classrooms where the guidelines proposed by the British Association of Teachers of the Deaf (BATOD) [49] ($T \leq 0.4$ s from 125 Hz to 4000 Hz) were applied, very favorable signal-to-noise ratios were obtained, following high levels of satisfaction by student and teachers.

According to the criterion of vocal comfort of Fig. 8, reverberation times slightly lower than 0.4 s can provide just 'acceptable' comfort (but not 'recommended' comfort). However, the recommended vocal comfort range was evaluated in laboratory experiments based on self-listening considerations, in quiet conditions. It is possible that, in classrooms with low reverberation times, as shown by the Essex study, noise levels are kept low and students are more engaged into the lesson, and the teacher becomes more satisfied about the overall acoustic situation.

Taking into account these studies, it can be concluded that, in classrooms for flexible teaching methods with no more than 40 students, reverberation times above 0.6 s (in conditions of full occupancy) or 0.7 s (in unoccupied furnished classrooms) should

be avoided and that volumes should be kept no larger than 210 m^3 (however, volumes can be larger for less students). Higher reverberation times would decrease speech intelligibility, which at the same time would disengage students from learning. Moreover, activity noise levels would increase due to the lower engagement, to the lesser absorption in the classroom, and to the Lombard effect. Whereas some guidelines, like the one proposed by the BATOD [49], stand for reverberation times shorter than 0.4 s based on speech intelligibility considerations, the present study suggests that reverberation times slightly higher (between 0.45 s and 0.5 s in occupied classrooms) might be optimal, and that lower reverberation times are not always better, based on vocal comfort considerations. Therefore, guidelines should include target reverberation times (or upper and lower limits) rather than maximum allowed reverberation times, and should include limitations in volume and number of students—at least regarding the necessity of having a dedicated study with geometrical considerations and distribution of absorbing material on the surfaces of the classroom, if designing larger classrooms for more students. However, in the case of classrooms especially designed for children with hearing impairment, acoustic requirements for students should have more weight than requirements for vocal comfort, the number of students in the classroom should be lower than 20, and reverberation times around 0.4 s or slightly lower should be sought, in accordance with the BATOD guidelines [49] and the recommendations of Nijs and Rychtarikova [41]. Nevertheless, reverberation times lower than 0.3 s might generate ‘overdamping’, i.e. an excessive attenuation of speech levels at the last rows [41].

In some European countries (see Table 2), like France or UK (for secondary schools), reverberation times up to 0.8 s in unoccupied spaces are allowed, which may be detrimental for speech intelligibility if, with students present, the reverberation time does not descend below 0.6 s. On the other hand, in Norway and Spain, reverberation times in empty furnished classrooms are limited to 0.5 s (0.4 s for the highest quality class in Norwegian classification scheme), which might be too restrictive and lead to low vocal

comfort levels. From the studied countries, only Germany and France set maximum and minimum limits to reverberation times. Regarding volume limitations, France includes the prescription of a dedicated study for classrooms larger than 250 m³, Spain for classrooms larger than 350 m³ and Germany gives special design guidelines for classrooms larger than 250 m³. According to the present findings, it is apparent that some guidelines should be adjusted: Denmark, Norway, Spain and UK should add minimum limits for their required reverberation times. France and UK (for secondary schools) should lower the maximum allowed reverberation time. Spain should lower the volume of 350 m³ for the need of doing a dedicated study. Moreover, Denmark, Norway and UK should take the volume of the classroom into consideration. At the same time, the slope of target reverberation time as a function of the volume of the classroom, as used in the German standard, might be excessive according to the vocal comfort and U_{50} criteria. Classification standards, such as the Norwegian NS8175:2012 should also be revised to consider what are the best quality standards and include target reverberation times instead of upper limits. According to the present findings, a limitation of the maximum reverberation time to 0.4 s in furnished but unoccupied classrooms for the highest quality classes is detrimental for vocal comfort.

Current European regulations differ in the methods they use to express the requirements in terms of reverberation time, volume, occupancy conditions and additional demands on the frequency dependency and validation of the design. With the existing knowledge in classroom acoustics, there is an excellent opportunity to work towards the harmonization of regulations among European countries.

Acknowledgments

The research on speaker-oriented classroom acoustics design guidelines was funded by the Swedish organization AFA Försäkring as a part of the project “Speakers’s comfort

and voice health in classrooms” and is currently funded by Research Foundation – Flanders. We thank the anonymous reviewers for their valuable comments that lead to significant improvements on the final version of this article.

Appendix: Validation of the prediction model for $DT_{40,ME}$

In order to validate the prediction model for $DT_{40,ME}$, a retrospective analysis of the measurements in 30 educational facilities was performed, the same ones used for validating the prediction model for ST_V [18].

The measurements were performed in 30 unoccupied but totally furnished rooms, with volumes in the range from 40.6 m^3 to 3600 m^3 and surface areas in the range from 52 m^2 to 1700 m^2 .

The reverberation time T was derived from measurements of room impulse responses using an omnidirectional source B& K type 4295 “Omnisource” at a height of 1.6 m and a microphone B& K type 4192 at a height of 1.2 m. The reverberation times were calculated from the decays between -5 and -25 dB of the backward integrated energy curves and the values for the octave bands of 500 Hz, 1 kHz and 2 kHz were averaged.

The OBRIR was measured with a Head and Torso Simulator B& K type 4128 positioned at a height of 1.5 m at two different locations in each room. The $DT_{40,ME}$ was derived for each of the octave bands between 125 Hz and 4 kHz, and was averaged for the two measurement positions and the two ears.

Figure 14 shows the predictions of the $DT_{40,ME}$ model — evaluated using the measured V , S , and T_{20} values — versus the measured values of $DT_{40,ME}$ in octave bands. In each frequency band, a regression line of the type $\overline{DT}_{40,ME,pred} = a \cdot DT_{40,ME,meas} + b$ is calculated, where $\overline{DT}_{40,ME,pred}$ is the regressor for the predicted values of $DT_{40,ME}$ (represented as $DT_{40,ME,pred}$), $DT_{40,ME,meas}$ are the measured values, and a and b are the coefficients of the regression line. Ideally, a perfect model would result if the pre-

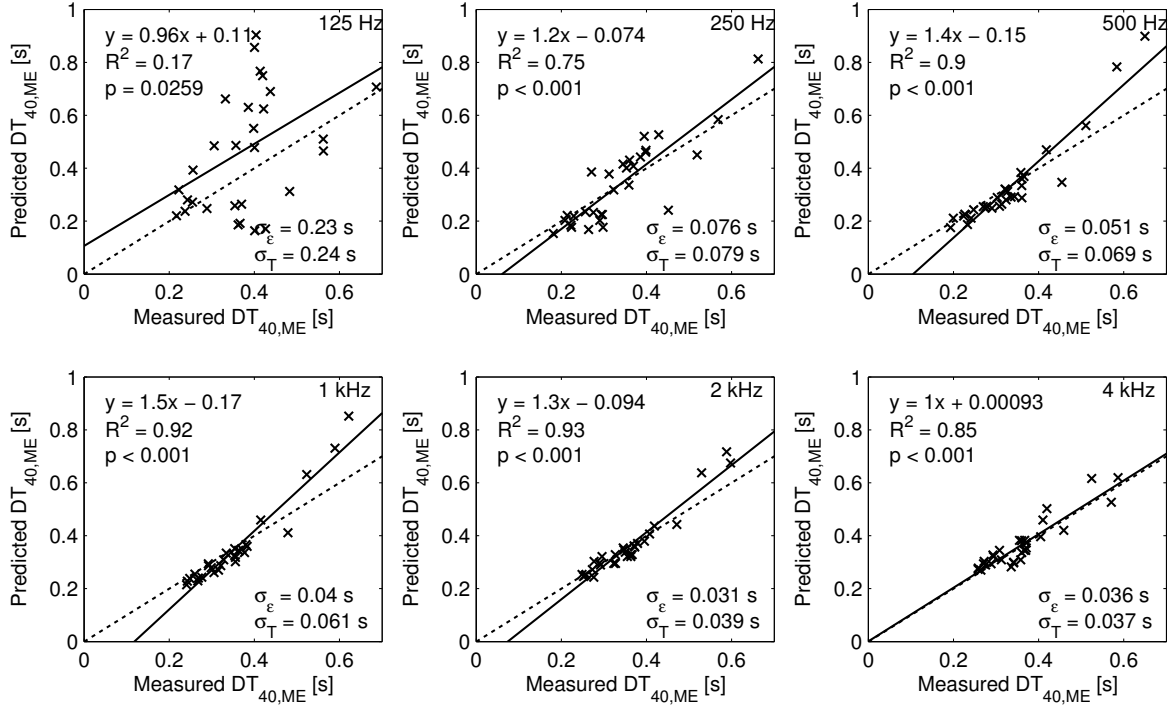


Figure 14: Predicted versus measured values of decay time $DT_{40,ME}$ in frequency bands. The solid lines show the regression lines for the predictions and the dotted lines indicate the unbiased prediction lines.

dicted and the measured values were equal ($DT_{40,ME,pred} = DT_{40,ME,meas}$). An unbiased model would result if $a = 1$ and $b = 0$, i.e. $\overline{DT}_{40,ME,pred} = DT_{40,ME,meas}$.

The goodness of fit of the prediction model is evaluated with three parameters: a) the coefficient of determination R^2 of the linear regression model for the measured versus predicted values, b) the residual deviation σ_ϵ of the predicted values from this regression line, and c) the deviation σ_T of the predicted values from an unbiased prediction, which is a measure of the bias in the prediction.

$$\sigma_\epsilon^2 = \frac{1}{N-2} \sum_{i=1}^N (DT_{40,ME,pred} - \overline{DT}_{40,ME,pred})^2 \text{ (s}^2\text{)}. \quad (25)$$

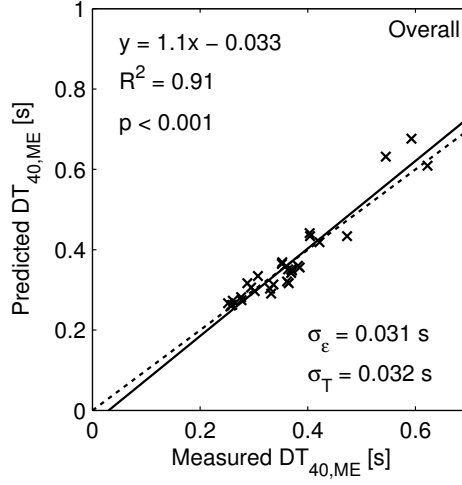


Figure 15: Predicted versus measured values of decay time $DT_{40,ME}$, averaged over the octave bands of 2 kHz and 4 kHz. The solid lines show the regression lines for the predictions and the dotted lines indicate the unbiased prediction lines.

$$\sigma_T^2 = \frac{1}{N} \sum_{i=1}^N (DT_{40,ME,pred} - DT_{40,ME,meas})^2 \text{ (s}^2\text{)}. \quad (26)$$

The most accurate predictions, as shown in Fig. 14, are at the octave bands of 2 kHz (highest R^2 , of 0.93) and 4 kHz, where there is negligible bias ($a \approx 1, b \approx 0$). Thus, the octave band values of 2 kHz and 4 kHz are averaged and the predicted versus measured values of decay time $DT_{40,ME}$ are shown in Fig. 15. $DT_{40,ME}$ averaged over 2 kHz and 4 kHz presents a low bias ($a = 1.1, b = 0.033$), a high coefficient of determination ($R^2 = 0.91$) and a standard deviation of the prediction error of 0.03 s.

References

- [1] D. Pelegrin-Garcia and J. Brunskog, “Speakers’ comfort and voice level variation in classrooms: Laboratory research,” *J. Acoust. Soc. Am.*, vol. 132, pp. 249–260, 2012.

- [2] H. Traunmüller and A. Eriksson, “Acoustic effects of variation in vocal effort by men, women, and children,” *J.Acoust.Soc.Am.*, vol. 107, pp. 3438–3451, 2000.
- [3] H. Schönwälder, J. Berndt, F. Ströver, and G. Tiesler, “Belastung und beanspruchung von lehrerinnen und lehrern (strain and stress of teachers),” Tech. Rep. Fb 989, Federal Institute for Occupational Safety and Health (BAuA), Bremerhaven: Wirtschaftsverlag NW, 2003.
- [4] J. Kristiansen, S. Lund, P. Møberg Nielsen, R. Persson, and H. Shibuya, “Determinants of noise annoyance in teachers from schools with different classroom reverberation times,” *J.Environ.Psychol.*, vol. 31, pp. 383–392, 2011.
- [5] J. Bradley, R. Reich, and S. Norcross, “On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility,” *J.Acoust.Soc.Am.*, vol. 106, pp. 1820–1828, 1999.
- [6] J. Bradley and H. Sato, “The intelligibility of speech in elementary school classrooms,” *J.Acoust.Soc.Am.*, vol. 123, pp. 2078–2086, 2008.
- [7] A. Astolfi, P. Bottalico, and G. Barbato, “Subjective and objective speech intelligibility investigations in primary school classrooms,” *J.Acoust.Soc.Am.*, vol. 131, pp. 247–257, 2012.
- [8] J. Bradley, “Speech intelligibility studies in classrooms,” *J.Acoust.Soc.Am.*, vol. 80, pp. 846–854, 1986.
- [9] T. Houtgast, “The effect of ambient noise on speech intelligibility in classrooms,” *Appl. Acoust.*, vol. 14, pp. 15–25, 1981.
- [10] J. Brunskog, V. Lyberg-Åhlander, A. Löfqvist, D. Pelegrin-Garcia, and R. Rydell, *Final report of the project “Speakers comfort and voice disorders in classrooms”*.

Lund, Sweden: Sound Environment Center at Lund University, 2011. ISSN 1653-9354.

- [11] V. Lyberg-Åhlander, R. Rydell, and A. Löfqvist, “Speaker’s comfort in teaching environments: Voice problems in Swedish teaching staff,” *J. Voice*, vol. 25, pp. 430–440, 2011.
- [12] J. Brunskog, A. Gade, G. P. Ballester, and L. R. Calbo, “Increase in voice level and speaker comfort in lecture rooms,” *J. Acoust. Soc. Am.*, vol. 125, no. 4, pp. 2072–2082, 2009.
- [13] D. Pelegrin-Garcia, “Comment on “Increase in voice level and speaker comfort in lecture rooms” [J. Acoust. Soc. Am. 125, 2072–2082 (2009)],” *J. Acoust. Soc. Am.*, vol. 129, pp. 1161–1164, 2011.
- [14] H. Lane and B. Tranel, “The Lombard sign and the role of hearing in speech,” *J. Speech Lang. Hear. Res.*, vol. 14, no. 4, pp. 677–709, 1971.
- [15] J. A. T. Whitlock and G. Dodd, “Speech intelligibility in classrooms: Specific acoustical needs for primary school children,” *Build. Acoust.*, vol. 15, pp. 35–47, 2008.
- [16] J. Rindel, “Verbal communication and noise in eating establishments,” *Applied Acoustics*, vol. 71, pp. 1156–1161, 2010.
- [17] M. Hodgson, R. Rempel, and S. Kennedy, “Measurement and prediction of typical speech and background-noise levels in university classrooms during lectures,” *J. Acoust. Soc. Am.*, vol. 105, pp. 226–233, 1999.
- [18] D. Pelegrín-García, J. Brunskog, V. Lyberg-Åhlander, and A. Löfqvist, “Measurement and prediction of voice support and room gain in school classrooms,” *J. Acoust. Soc. Am.*, vol. 131, pp. 194–204, 2012.

- [19] B. Rasmussen, J. Brunskog, and D. Hoffmeyer, “Reverberation time in classrooms—Comparison of regulations and classification criteria in the Nordic countries,” in *Proceedings of Joint Baltic-Nordic Acoustics Meeting BNAM2012*, (Odense, Denmark), pp. 1–6, 2012.
- [20] Deutsche Institut für Normung e.V., *DIN 18041:2004 Acoustical quality in small to medium-sized rooms*. Berlin, Germany, 2004.
- [21] Danish Enterprise and Construction Authority, Danish Ministry of Economic and Business Affairs, *Building Regulations 2010*. Copenhagen, Denmark, 2010. http://erhvervsstyrelsen.dk/file/155699/BR10_ENGLISH.pdf (last accessed 19 June 2013).
- [22] D. Hoffmeyer, *Lydforhold i undervisnings- og daginstitutionsbygninger - Lydbestemmelser og anbefalinger (Acoustic conditions in educational and day care buildings—Regulations and recommendations)*. SBI Guideline 218. Aalborg University, Hørsholm, Denmark, 2010.
- [23] French Council of State, *Code de la construction et de l’habitation (French Building Code)*. Paris, France, 2012. Version of 10th September 2012.
- [24] French Ministry of Ecology and Sustainable Development, *Arrêté du 25 avril 2003 relatif à la limitation du bruit dans les établissements d’enseignement (decree of 25 april 2003 relative to the noise limitation in educational settings)*. Paris, France, 2003.
- [25] French Ministry of Ecology and Sustainable Development, *Circulaire du 25 avril 2003 relative à l’application de la réglementation acoustique des bâtiments autres que d’habitation (Circular of 25 april 2003 relative to the application of Acoustique regulation for buildings other than apartment buildings)*. Paris, France, 2003.

- [26] National Office of Building Technology and Administration, *Teknisk Forskrift 2010 (TEK'10) Technical Regulations under the Planning and Building Act 1997*. Norway, 2010.
- [27] Norge Standard, *NS 8175:2012—Lydforhold i bygninger, Lydklassifisering for ulike bygningstyper (Acoustic conditions in buildings - Sound classification of various types of buildings)*. Norway, 2012.
- [28] Spanish Housing Ministry, *Codigo Técnico de la Edificación (Technical Building Code). RD 314/2006*. Madrid, Spain, 2006.
- [29] Spanish Housing Ministry, *Documento Básico HR - Protección Frente al Ruido (Basic Document HR - Protection against Noise)*. Madrid, Spain, 2009.
- [30] HM Government, *The Building Regulations 2010. Approved document E: Resistance to the passage of sound. 2003 edition incorporating 2004 and 2010 amendments*. London, UK, 2010.
- [31] Department for Education and Skills, *Building Bulletin 93. Acoustic Design of Schools: A Design Guide*. London, UK, 2003.
- [32] Austrian Standards Institute, *ÖNORM B 8115-3—Schallschutz und Raumakustik im Hochbau - Teil 3: Raumakustik (Sound insulation and architectural acoustics in building construction - Part 3: Architectural acoustics)*. Viena, Austria, 2005.
- [33] Swiss Society of Engineers and Architects, *SIA 181:2006—Schallschutz in Hochbau / Protection contre le bruit dans le bâtiment / La protezione dal rumore nelle costruzioni edilizie (Protection against noise inside buildings)*. Zürich, Switzerland, 2006.
- [34] NBN - Bureau de Normalisation, *NBN S01-400-2: Acoustic requirements in school buildings*. Brussels, Belgium, 2012.

- [35] Swedish Standards Institute, *SS 25268:2007–Byggakustik - Ljudklassning av utrymmen i byggnader - Vårdlokaler, undervisningslokaler, dag- och fritidshem, kontor och hotell (Acoustics - Sound classification of spaces in buildings - Institutional premises, rooms for education, preschools and leisure-time centres, rooms for office work and hotels)*. Sweden, 2007.
- [36] Icelandic Standards,, *IST 45:2011–Hljóðvist—Flokkun íbúðarog atvinnuhúsnæðis (Acoustic conditions in buildings - Sound classification of various types of buildings)*. Iceland, 2011.
- [37] D. Pelegrín-García, B. Smits, J. Brunskog, and C.-H. Jeong, “Vocal effort with changing talker-to-listener distance in different acoustic environments.,” *J.Acoust.Soc.Am.*, vol. 129, pp. 1981–90, 2011.
- [38] D. Pelegrín-García, O. Fuentes-Mendizabal, J. Brunskog, and C.-H. Jeong, “Equal autophonic level curves under different room acoustics conditions,” *J.Acoust.Soc.Am.*, vol. 130, pp. 228–238, 2011.
- [39] D. Pelegrin-Garcia, V. Lyberg-Åhlander, R. Rydell, J. Brunskog, and A. Löfqvist, “Influence of Classroom Acoustics on the Voice Levels of Teachers With and Without Voice Problems: A Field Study,” *Proceedings of Meetings on Acoustics*, vol. 11, pp. 1–9, 2010.
- [40] H. Lazarus, “Prediction of Verbal Communication in Noise-A Review: Part 1,” *Appl. Acoust.*, vol. 19, no. 6, pp. 439–463, 1986.
- [41] L. Nijs and M. Rychtarikova, “Calculating the optimum reverberation time and absorption coefficient for good speech intelligibility in classroom design using U50,” *Acta Acust. united Ac.*, vol. 97, no. 1, pp. 93–102, 2011.
- [42] M. Barron, *Auditorium acoustics and architectural design*, ch. Objective measures for speech auditoria, pp. 425–429. London: Taylor & Francis, 1993.

- [43] H. Sato and J. Bradley, “Evaluation of acoustical conditions for speech communication in working elementary school classrooms,” *J.Acoust.Soc.Am.*, vol. 123, pp. 2064–2077, 2008.
- [44] F. L. Rivera-Batiz and L. Marti, “A school system at risk: A study of the consequences of overcrowding in new york city public schools,” *IUME Research Report*, vol. 95, 1995.
- [45] M. Picard and J. Bradley, “Revisiting speech interference in classrooms,” *Audiology*, vol. 40, pp. 221–244, 2001.
- [46] B. Shield and J. Dockrell, “The effects of noise on children at school: a review,” *Building Acoustics*, vol. 10, no. 2, pp. 97–116, 2003.
- [47] American National Standards Institute, *ANSI/ASA S3.5-1997 (R2012)–American National Standard Methods for Calculation of the Speech Intelligibility Index*. New York, NY, 2012.
- [48] D. Canning and A. James, *The Essex Study. Optimised classroom acoustics for all*. St Albans, UK: The Association of Noise Consultants, 2012.
- [49] The British Association of Teachers of the Deaf (BATOD), *Class-room Acoustics–Recommended Standards*. High Wycombe, UK, 2001.
<http://www.batod.org.uk/content/resources/audiology/magazine/acoustic-standards.pdf> (last accessed 20 June 2013).